EXPERT REVIEWS

The use of reinforcement learning algorithms to meet the challenges of an artificial pancreas

Expert Rev. Med. Devices 10(5), 661–673 (2013)

Melanie K Bothe[1], Luke Dickens[2], Katrin Reichel[1,3], Arn Tellmann[1], Björn Ellger[4], Martin Westphal[1,4] and Ahmed A Faisal*[2,3,5]

[1]*Fresenius Kabi Deutschland GmbH, Else-Kröner-Strasse 1, 61352 Bad Homburg, Germany*
[2]*Department of Computing, Huxley Building, Imperial College London, London SW7 2AZ, UK*
[3]*Department of Bioengineering, Royal School of Mines Building, Imperial College London, London SW7 2AZ, UK*
[4]*Department of Anesthesiology, Intensive Care and Pain Medicine, University Hospital of Muenster, 48149 Muenster, Germany*
[5]*MRC Clinical Sciences Center, Hammersmith Hospital Campus, London W12 0NN, UK*
**Author for correspondence:*
Tel.: +44 (0)781-5863146
a.faisal@imperial.ac.uk

Blood glucose control, for example, in diabetes mellitus or severe illness, requires strict adherence to a protocol of food, insulin administration and exercise personalized to each patient. An artificial pancreas for automated treatment could boost quality of glucose control and patients' independence. The components required for an artificial pancreas are: i) continuous glucose monitoring (CGM), ii) smart controllers and iii) insulin pumps delivering the optimal amount of insulin. In recent years, medical devices for CGM and insulin administration have undergone rapid progression and are now commercially available. Yet, clinically available devices still require regular patients' or caregivers' attention as they operate in open-loop control with frequent user intervention. Dosage-calculating algorithms are currently being studied in intensive care patients [1], for short overnight control to supplement conventional insulin delivery [2], and for short periods where patients rest and follow a prescribed food regime [3]. Fully automated algorithms that can respond to the varying activity levels seen in outpatients, with unpredictable and unreported food intake, and which provide the necessary personalized control for individuals is currently beyond the state-of-the-art. Here, we review and discuss reinforcement learning algorithms, controlling insulin in a closed-loop to provide individual insulin dosing regimens that are reactive to the immediate needs of the patient.

Keywords: artificial pancreas • continuous glucose monitoring • diabetes mellitus • machine learning • personalized medicine • reinforcement learning

Maintaining normoglycemia is one of the major challenges in the treatment of patients with diabetes mellitus. Treatment of hyperglycemia with insulin may lead to hypoglycemia that, in turn, may contribute to clinically relevant complications. This implies that calculation of precise insulin dosages is critical, must be individually adapted and should be reactive to the patient's glucose concentration. Worldwide, more than 371 million people have diabetes mellitus [4], and its management can involve both constant glucose monitoring and insulin dosing, seriously affecting quality of life. This has led to intensive research concerning the development of an artificial pancreas since the 1970s [5].

There are three components of an artificial pancreas: i) continuous glucose monitoring (CGM) using an implanted sensor, ii) an insulin pump delivering insulin and iii) an algorithm calculating the correct dose of insulin to be applied. In this article, we briefly describe the major challenges for development of an artificial pancreas system and discuss the application of machine learning algorithms as a potential approach to increase the efficiency of the system in terms of versatility and safety. Since the development of the first artificial pancreas system [5], major improvements have been made, but the system still needs development before it can be routinely used in clinical practice and everyday life. To date, one of the major limitations of the successful use of automated dosing in clinical, as well as outpatient, settings is the demand for an adaptive algorithm that individualizes the artificial pancreas to the special needs of each patient. Independent of the control model used, the components

 10.1586/17434440.2013.827515 ISSN 1743-4440

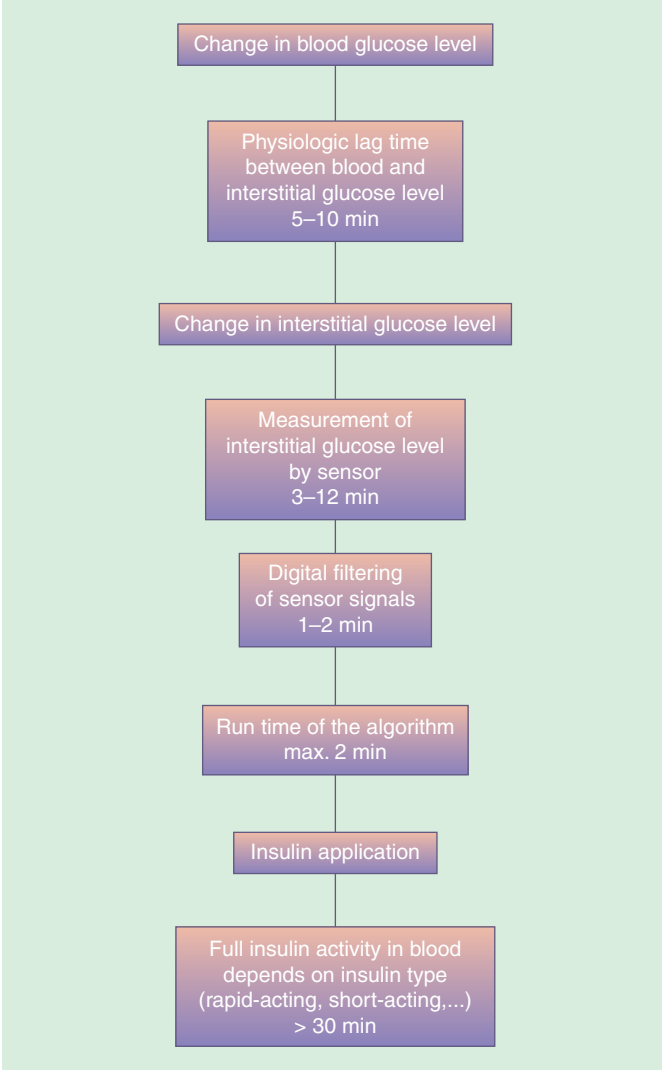

Figure 1. Stages of the time delay between blood glucose concentration and maximum blood insulin concentration leading to a lag time of approximately 1 h.

of an artificial pancreas system all bear their own challenges. These are either derived from technical aspects or based on the physiology of glucose regulation and have to be kept in mind during the development of control algorithms for insulin dose calculation.

Continuous glucose monitoring

Measuring the blood glucose concentration provides the minimal requirement for calculation of the insulin dose. Patients with type 1 diabetes are recommended to check their blood glucose concentrations at least three times a day [6]. More frequent glucose measurements allow for the overall trend of the blood glucose concentration to be estimated, as opposed to isolated measurements with no information on whether blood glucose concentrations are increasing or decreasing [7]. As finger pricking can be painful, a less invasive method enabling the patient to perform more frequent measurements is highly desirable.

Among various CGM sensors (for review on CGM sensors please refer to [7]), subcutaneously implantable sensors are most comfortable for the patient. Due to an evaluation roadmap for CGM sensors presented by the Clinical and Laboratory Standards Institute [8], the most important challenges for CGM devices to date relate to the accuracy of the measurement, as well as to the real-time assessment.

The accuracy of subcutaneous glucose measurement devices has been subject of a long-lasting debate. In such devices, the sensor detects glucose in the interstitial fluid during its diffusion between the capillary and the target cell [9]. Under steady-state conditions, interstitial glucose concentrations have been shown to be similar, but not precisely equal to, venous blood glucose concentrations in healthy individuals or animals [10–12]. Rapid changes in blood glucose concentrations have been reported to affect the accuracy of the interstitial glucose sensing, namely causing the sensor to report glucose concentrations below their actual values [13,14]. Implanted intravenous glucose sensors, which would provide similar comfort, as well as faster and more accurate blood glucose measurements, are currently under investigation [15,16]. However, to date subcutaneous measurement is still preferred due to lower risk of thrombosis and intravascular infection.

Real-time assessment of subcutaneous glucose measurement devices describes the lag between measurement of the glucose concentration by the sensor and the time at which the blood concentration of insulin, delivered in response, reaches its maximum. A large delay reduces the ability of the system to respond to glucose concentrations in real-time and therefore its performance. Figure 1 summarizes the sequence of events, from a change in blood glucose in the body to the maximum effect of the insulin administered in response [17]. First, changes in blood glucose concentrations are mirrored by the interstitial blood glucose concentrations after a 5–10 min delay [18,19]. Measurement of those interstitial glucose concentrations is commonly performed either by electrochemical sensors [20] or microdialysis techniques [21], which both take another 3–12 min [18]. Next, the digital filtering of the glucose measurement can take another 1–2 min, and is required to compensate for background noise. At this point, the algorithm will take some time to calculate the correct dose, but it is expected to be comparatively rapid. Finally, after insulin application, there is a delay before insulin becomes fully active in the blood. The latter depends on the type of insulin analog used, the total insulin dose and the individual pharmacokinetic parameters of the patient. Taken together, this can introduce a time delay between changes in blood glucose and insulin effect of up to 1 h, which any dosage calculation will have to accommodate.

Drug delivery: insulin administration

Since the first studies on insulin treatment of patients with diabetes [22,23], which have led to the Nobel prize for Frederick Banting and John Macleod, insulin delivery devices have seen a number of development phases that have improved their performance and ease of use. For example, the time delay between

application of insulin and the maximum plasma insulin concentration mentioned above have already been shortened by the availability of rapid-acting insulins. Furthermore, other administration routes also bear the potential to decrease the time until the maximum effect of insulin occurs.

Intravenous insulin application, which would enable the fastest insulin effect, exhibits certain limitations due to catheter complications [24]. Intraperitoneal insulin is difficult to administer [25], and the availability of intraperitoneal insulin devices is currently limited. Inhalation might provide a novel application route for insulin, and suitable medical devices have been approved by the US FDA in 2008 [26]. But the bioavailability of inhaled insulin is less than in a subcutaneous application and is extremely variable in smokers or patients with a cold [27]. Pharmaceuticals for oral delivery of insulin are currently under development [28,29] but are far from routine clinical use.

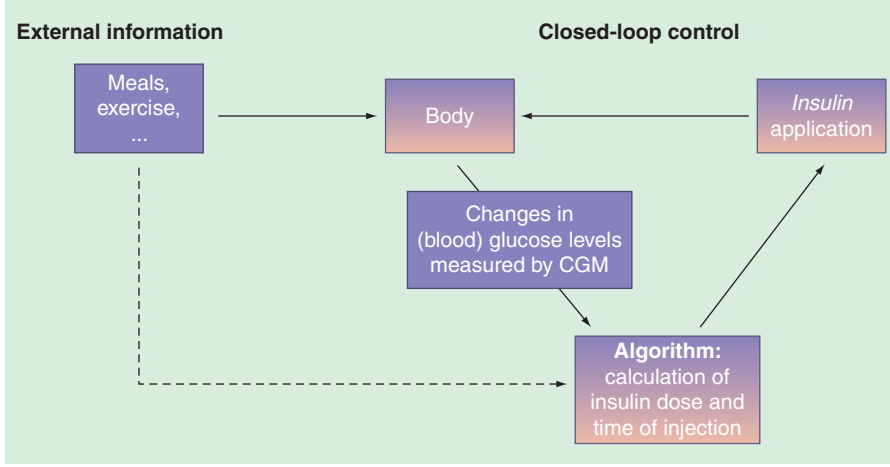

Figure 2. Closed-loop control model for insulin delivery with (hybrid model, dotted line) and without external information. In the latter case, the algorithm reacts directly on the changes in glucose concentrations evoked by meals or exercise without getting external information.

Wide availability and ease of management are the major advantages of subcutaneous insulin administration [24]. This currently makes the subcutaneous application route the most appropriate for routine injections and thus also for an artificial pancreas. When CGM is also performed subcutaneously, the system is commonly referred to as subcutaneous-subcutaneous (sc-sc) systems.

Dosage decisions: the controller is key

The control algorithm represents the key component of the artificial pancreas, because it provides the link between changes in blood glucose concentrations and the respective treatment response, that is, insulin delivery. In contrast to conventional insulin treatment with pre-programmed pumps, an artificial pancreas aims at modulation of insulin delivery in intervals close to real-time, in response to blood glucose levels, and facilitated by a control algorithm.

Contemporary algorithms forming the 'brains' of any automated drug delivery device attempt at dealing with both the challenges derived from the limits of CGM and insulin as well as from the glucose regulatory system itself. In the next section, different types of control algorithms as well as challenges of the glucose regulatory system for such algorithms are described and discussed.

Control algorithms & challenges

In standard diabetes treatment, the patient receives a subcutaneous injection of slow-acting insulin to provide the basal insulin requirement. Additional insulin doses of rapid-acting insulin are calculated based on the patient's knowledge of a meal size, the patients experience, the insulin sensitivity and actual blood glucose concentrations measured indirectly, for example, by a subcutaneous CGM system. Insulin boluses are preferably delivered by an insulin pump to avoid repeated injections. As a key performance indicator, an average plasma glucose concentration of glycosylated hemoglobin (HbA_{1C}) values <7% was recommended by the American Diabetes Association [6] and has been shown to reduce the development and progression of microvascular and cardiovascular complications by 76% [30].

Open-loop vs closed-loop control

Models that calculate an insulin dosage based on blood glucose concentrations and external (meal) information have been called examples of open-loop control [27]. Despite the advantages of the intended avoidance of hyper- and hypoglycemic excursions by use of an artificial pancreas system, any open-loop insulin control mechanism requires the patient to live a more or less predictable lifestyle. By contrast, closed-loop control models include the benefit to finally reduce the patients' need to plan each day with regard to their illness, thereby improving quality of life (for review about open- and closed-loop models please refer to [31]). Closed-loop models can be further distinguished as either fully closed-loop models or hybrid models. In a fully closed-loop model, decisions for insulin dosing are exclusively based on parameters measured in the patient's body, for example, blood glucose concentrations [27] without knowledge of external information like food or exercise. In brief, changes in blood glucose concentrations would be measured and, based on these changes, the respective insulin dosage calculated and applied. Administration of insulin then affects the blood glucose concentration. Based on this feedback from the blood glucose concentration, the required concentration of insulin is again calculated and the next dose may be adjusted (Figure 2). Such a system is reactive, meaning no anticipation and preemptive dosing based on external information is possible (i.e., did the patient enter an ice cream shop). However, the closed-loop algorithms could factor in prior experience of glucose dynamics. A second major challenge for fully closed-loop models is to accommodate situations

in which the blood glucose concentration changes rapidly, such as after meals or exercise. Adequate and timely reactions of the artificial pancreas on these rapid changes are hindered by a large time delay due to both glucose measurement as well as insulin action. One would assume that the real pancreas works in a closed-loop system without this knowledge as well, however, recent work suggests that there's a tighter link between brain and pancreatic function [32,33].

Hybrid control

Hybrid models using both closed-loop control and external information, for example, of meal size (Figure 2), have therefore been proposed. In a recent study, Weinzimer *et al.* compared fully closed-loop control with their hybrid or extended closed-loop control, in which an additional premeal bolus was used [34]. The hybrid model reduced hyperglycemia after meals without inducing hypoglycemia [34], and the same result was seen in a similar model tested by another group in hospital settings [35]. In both studies, patients were already in excellent glycemic control, and thus hypoglycemia was neither observed with the hybrid, nor the fully closed-loop control model [34,35]. However, these hybrid models still need to be tested for control of potentially hypoglycemic conditions and in their present form need explicit user input and action. Hybrid models may have a physiological basis, as recent data suggest hypothalamic–pancreatic connectivity (from central to autonomous nervous system [32,33]) which could convey information about planned meal intake and could potentially therefore be closer to the type of information available to the pancreas, than closed loop models. However, the brain's knowledge that 'I am queuing for an ice cream' does not inform the pancreas about the exact composition of the meal. Therefore, it has to be tested whether the hybrid model might in the future be replaceable by a fully closed-loop model to reduce the required frequency of user interventions.

Despite the good performance of hybrid closed-loop model with external information, the overall aim is to develop a fully closed-loop model to obviate the need for diabetes patients to lead a scheduled lifestyle in regard to meals and exercise. One way to achieve this is to develop an algorithm that automatically detects meals by checking the blood glucose curve and either advises the patient or the automated insulin pump to apply an insulin bolus. A first attempt to develop an analytic model that, at least partially, anticipates fluctuations in glycemia, was performed by development of algorithms for meal detection or meal size estimation [36,37]. When compared *in silico* with closed-loop control without information on meals, the meal size estimation algorithm improved the time spent in normoglycemic range and even reduced the HbA_{1C} from 7.15% (treatment without algorithm) to 6.43% (treatment with algorithm) in adolescents and from 6.69 to 6.23% in adults [37]. However, the average detection time of the onset of a meal was at least 29 or 31 min, respectively [36,37]. Given the additional delay of insulin action, this might be too late for appropriate insulin delivery using current insulin formulations. Furthermore, the amount of false-positive (6.75%) and false-negative (18%) detections [37] increases the risk of incorrect insulin administration and has to be improved to facilitate a safe application in patients. Long-term studies should be performed in various settings to detect whether this addition of complexity to the glucose control system is worth the effort. Meal detection is only one out of many challenges for closed-loop systems of insulin delivery control.

Traditional algorithms for control: PID, MPC, Fuzzy

The development of algorithms for closed-loop calculation of insulin dosage is intensively investigated. The major candidates for such algorithms proposed in recent years use proportional integrative derivate (PID [35,38–40]) methods or model predictive control (MPC [1,37,41–56]). The PID systems consist of three components: the proportional (in case of diabetes the difference between the actual glucose concentration and the desired glucose concentration – the error), the integral (accumulation of past errors over time) and the derivative (the rate of change of these errors) [57]. In short, the PID algorithm estimates the required control (in case of diabetes the required delivery of insulin) based on a weighted sum of PID terms, in order to minimize these errors and so bring the system to the desired glucose concentration [57]. Current MPC systems require a model (typically a dynamical systems model) that can predict future glucose concentrations given known values for current glucose, insulin delivery and food intake. Such control then calculates the appropriate insulin infusion rate by minimizing the difference between the model-predicted glucose concentration and the target glucose concentrations over a prediction time-window [2,58]. The duration of this time-window is chosen as the time in which the bulk of the effect is seen from the insulin or insulin analog used.

Conventional MPC algorithms suffer a number of issues that would need to be addressed for their safe application in outpatient use. First, MPC assumes a perfect model and control calculations are based on this assumption. Imperfections in the model lead to differences between the expected and actual evolution of the system, and this means that the control must be recalibrated at each step. This brings an online computational expense, and more importantly there is no explicit information about how uncertain the model is (see section 'How to control in the face of uncertainty?'). Second, conventional MPC operates on a purely dynamical systems description model [2,58], and external disruptions to the system due to, for instance, meal intake or physical activity, which are not captured by the model, can only be reacted to after the event; MPC cannot compensate for the risk of these events in advance.

Recent advances in MPC control attempt to compensate for the first of these shortcomings. For instance in [48], the authors adapt the MPC algorithm to target a range of states instead of a single target state, and the approach proposed in [51] allows the aggressiveness of the control to be adapted in order to trade-off the rapidity of a return to normoglycemia against the amount of insulin used (and hence the risk of overshoot into

hypoglycemia); a trade-off made necessary by the imperfections of the model. These approaches are not truly adaptive in the sense that they shape their models in response to the available data. Instead, they are heuristic tuning approaches using a static model, and deviations of future state from what is predicted by the model are, just like conventional MPC, responded to in subsequent time-steps by updated responses to this state (again based on the imperfect model). To our knowledge, there are no existing MPC methods that address the second issue. Similar approaches using optimal control instead of MPC have also been suggested in the context of the artificial pancreas [39]. To the authors' knowledge, appropriate methods have yet to be realized, and these would suffer similarly in terms of reliance on appropriate models.

Fuzzy logic approaches have also been proposed for the control of the artificial pancreas, for example, see [3,59]. Fuzzy logic-based control approaches for the artificial pancreas use a system of if-then-else statements to determine when to apply insulin and the associated dose. These methods are not based on crisp empirical models of a patient's metabolism, but are sets of binary conditions and must be developed in collaboration with experts, for example, caregivers [3]. For this reason, each condition-response rule must be explicitly encoded into the policy, are subject to human error and omission and they cannot readily incorporate knowledge from existing biological models, so there can be no theoretically grounded performance guarantees.

In summary, the PID and fuzzy logic approaches are purely reactive and lack the theoretical underpinnings of a model, whereas MPC algorithms represent a more proactive approach, but require a good model of the dynamics and have a limited capacity to compensate for an imperfect and/or changing model. Ultimately, all existing MPC approaches for the artificial pancreas are based on a model that is defined at the level of the pancreas, but the glucose-insulin regulatory system does not represent the whole picture (see section 'How to control in the face of uncertainty?').

To date, there is only one closed-loop algorithm commercially available, that is, the B. Braun Space GlucoseControl (B. Braun, Melsungen, Germany). This algorithm is based on MPC [1,60] and is provided for use in insulin treatment of critically ill patients in intensive care units (ICU). Such patients often develop peripheral insulin resistance and relative insulin deficiency with resulting hyperglycemia [61]. Among others, this endocrine paradigm leads to increased gluconeogenesis from the body's stores and reduced glucose uptake and utilization. The former notion that resulting hyperglycemia would redistribute glucose toward organs that rely on glucose as fuel and, consecutively, improve the chance to survive was disapproved by evidence [62]. Indeed, dysregulations of glycemia are associated with a negative outcome. However, bringing glucose back to normoglycemia in the critically ill by insulin infusion has shown to be a double-edged sword, since hypoglycemia and fluctuations of glycemia offset beneficial effects of glucose control when the target range is set too low [61]. Therefore, critical care societies recommend controlling glycemia below 145 or 180 mg/dl, respectively [6]. Glycemic control by the one commercial algorithm initially required an hourly measurement of glucose [63], but recent adaptation of the algorithm has derestricted this constraint [64]. This algorithm is currently only used for critically ill patients and is exclusively applicable in the special setting of intensive care. Though integrated closed-loop glucose control is very promising in clinical settings [65], algorithms for home patients might raise the concern of missing control by human sense. This can be overcome by including an optional user-check of the insulin dosage values calculated by the algorithm, for example, by providing the algorithm as a smartphone application as recently reported by Cobelli *et al.* [66]. Despite the potential inclusion of a user-check, further improvements to the accuracy and robustness of an insulin calculating algorithm for independent outpatient use are highly desirable.

How to control in the face of uncertainty?

Uncertainty is a general concept in control and is the result of noise (randomness), ambiguity about the true state of variable and any other imperfection in the controller's predictive capabilities. Any control algorithm (and likewise the pancreas) has to operate in the face of uncertainty about: i) the physiological state, including the glucose level in the blood and the amount of ingested sugars in the gut, etc.; ii) variability in physiological processes across individuals (and within individuals during the course of the day); iii) the true dynamics of the model, as the interdependency between observed and unobserved variables may only be partly understood; iv) the evolution of the biological system, even if state and model is known precisely, system noise leads to uncertain outcomes; v) exogenous future events, such as food intake and physical activity which induce significant disruptions to the dynamics and vi) time delays in sensing state, this increases uncertainty about the current state meaning decisions are based on past information that may no longer be valid.

There are four major challenges to control the glucose system with an artificial pancreas as optimal effectiveness and performance require i) accurate sensor readings, with minimal time delays in ii) reading out physiological variables (sensing) and iii) acting on these (drug delivery action and effect). Finally, as every patient is different and has a different metabolism and lifestyle that can change with time, there is a need for iv) genuinely personalized medical treatment that adapts to the individual and their changing needs.

Of these challenges, the inter- and intraindividual variability of the diabetes patients most starkly raises the need for individual algorithms reacting on the special glucose-insulin pattern of each single patient. The differences in severity of the disease as well as in insulin sensitivity between patients and even during the life of a single patient require a continuous adaptation of the algorithm to the patient. Furthermore, individual meal uptake has a highly variable impact on the glucose concentration in the blood depending on the composition of the meal. The effect of situations like exercise, different levels of stress or illness on glucose homeostasis varies as well. These examples show that an algorithm for an artificial pancreas has to react

very quickly to rapidly changing situations influenced by abundant parameters to provide a sensible calculation of insulin.

For these reasons, we need to use techniques that adapt to the system as a whole, not least because the mechanism underlying the control of the biological pancreas is much more involved than was previously thought [67]. The algorithm has to learn models that are rich enough to describe the interplay between food intake, physical activity and insulin dynamics in the wild. We argue, that machine learning in general, and reinforcement learning (RL) in particular, provides the tools to describe and control such systems in an expressive and adaptive way.

Future trends

In order to develop an algorithm appropriate for an outpatient's artificial pancreas, a number of issues still need to be addressed including the response to meals, exercise, stress and sleep. To adequately deal with these situations, the algorithm either has to be explicitly developed to identify and respond to each situation separately, or its overall flexibility will have to improve, that is, by more closely mimicking the physiological function of a working human pancreas. The former approach is not promising, and so in the following section we will discuss the potential for improving the flexibility of such an algorithm.

Adding diagnostic indicators

Glucose homeostasis is maintained by various parameters including glucagon, epinephrine, insulin and others [68]. Some of these parameters may be useful indicators of food intake or stress levels for closed-loop control systems, obviating the need for external information. Incretins, for example, could be a useful indicator of food intake. Upon ingestion of a meal, incretins such as the glucagon-like peptide 1 (GLP-1) and gastric inhibitory peptide (GIP) promote the first phase secretion of insulin in proportion of the glucose content of the meal [68,69]. Measurement of active GLP-1 after meals could therefore indicate that glucose concentrations will soon be elevated and insulin dosing could be administered earlier, in anticipation of this. Unfortunately, GLP-1 is degraded by the dipeptidyl peptidase IV after approximately 2 min [69], making it difficult to detect. Present assays available for GLP-1 detection would have to be refined before routine measurement of GLP-1 could be used in this way.

After measuring a parameter like glucose or in the future GLP-1, the relationship between the parameter and the appropriate insulin dose response will have to be found. The first, original 'minimal model' of the insulin–glucose relationship was based on mathematical models of glucose and insulin kinetics only [70], subsuming the regulatory roles of many organ systems in these two compartments. Instead, current research uses models of physiologically based pharmacokinetics/pharmacodynamics. In these models, each organ system is treated as a separate compartment [68]. For example, the mixed meal model of Roy and Parker displays the absorption of the major compartments of a meal from the gut [71]. These models are currently used for simulation of diabetes and evaluating algorithms *in silico*. However, they can also be used to predict the pharmacodynamic response to hypothetical dosing strategies, and may, therefore, be incorporated into a future algorithm.

Beyond insulin: drugs regulating glucose homeostasis

Mimicking the physiological insulin pharmacokinetics in diabetes treatment is very demanding, especially for a dosage calculating algorithm. In a healthy pancreatic beta-cell, an increased blood glucose concentration and therefore the stimulation of the beta-cell by glucose induces a biphasic insulin release [72]. Due to the increment in plasma glucose concentrations, a rapid peak of insulin secretion is followed by a slowly increasing second phase of insulin secretion [72]. The peak phase of insulin secretion is due to pre-formed insulin stored in mature vesicles and is thought to suppress the hepatic glucose output [73]. The second phase requires new synthesis of proteins and increases slowly until the cell is adapted or the glucose stimulation ends. This biphasic profile of insulin secretion presents the researchers with the need of implicating the pharmacokinetics of different analogs of insulin into the calculations of the control algorithm. The time delay between application of insulin and the maximum plasma insulin concentration mentioned above have already been shortened by the availability of rapid-acting insulins (for review on insulin analogs see [74]). However, the fact that minimizing this time delay would markedly help to ensure normoglycemia in closed-loop models of artificial pancreas has raised the request for ultra-rapid-acting insulins. To achieve such a biphasic profile of insulin secretion, computational scientists need to include the pharmacokinetics of different analogs of insulin into the calculations of the control algorithm.

Another key player in glucose homeostasis is glucagon. Its role has been included in recent dynamic models and inclusion of such models in the artificial pancreas would more closely mimic the physiological pancreas reactions and thereby increase the flexibility and accuracy of the system. In a healthy pancreas, glucagon counters the effects of insulin, thus leading to elevation of blood glucose concentrations. Recently developed artificial pancreas systems apply pumps capable of both insulin and glucagon application [75]. Two independent research groups reported glucagon treatment to prevent and to reverse hypoglycemia in bihormonal closed-loop systems [75,76]. Minimizing glucagon dosage in two studies avoided side effects like nausea or gastrointestinal discomfort [75,76], though long-term studies remain to be conducted. Thus, glucagon treatment represents a promising option and should be considered in future development of control algorithms for artificial pancreas.

In addition, recent research has focused on the usability of new drugs regulating the glucose homeostasis. Amylin, for example, is a peptide hormone co-secreted with insulin by the pancreatic beta-cells which has similar functions to insulin [77,78]. Application of amylin was successfully tested in clinical trials [79] and even in closed-loop systems of combined insulin and amylin delivery [80]. In addition, the peptide GLP-1 could not only be measured for prediction of the correct insulin dose but may also be administered in addition to insulin [81,82]. However, exenatide, the first US FDA approved GLP-1 agonist, is still

assigned with a safety alert because it was associated with altered kidney function as well as hemorrhagic and necrotizing pancreatitis [83].

All these modifications mentioned might increase the flexibility of the artificial pancreas systems, but may also increase the complexity of the algorithms used. This issue might be addressed by using learning algorithms, which allow for dosing control to be flexibly optimized with respect to the biological system.

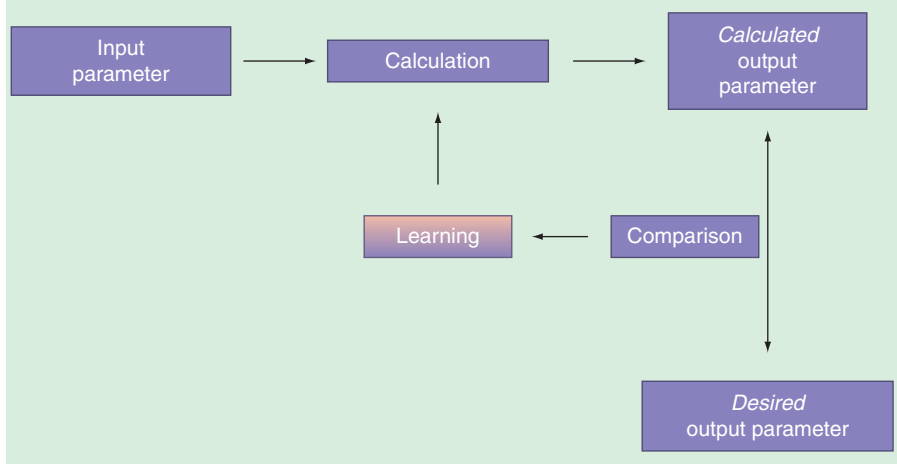

Figure 3. Principle of a learning algorithm. Comparison of the calculated output parameter with the desired output parameter leads to learning of the algorithm.

Personalized medicine through (machine) learning algorithms

The complexity of insulin delivery and the demanding goal of maintaining normoglycemia necessitate a complex, adaptive and flexible algorithm, which may be achieved with the use of machine learning techniques as shown in some approaches [84–90]. Appropriate machine learning algorithms are able to analyze training data, recognize complex patterns and on the basis of such patterns apply the knowledge to other data to predict their behavior [91]. The principle of a learning algorithm is depicted in Figure 3.

There are three general differences between traditional PID and MPC algorithms used for diabetes so far and machine learning approaches [91]. First, machine learning is based on recognition of patterns instead of implication of defined hypotheses. It improves the accuracy of the system, because it includes initially unidentified variables that might be overlooked by the traditional hypothesis-based systems [91]. A second advantage of machine learning approaches is that they consider interactions between variables instead of minimizing or ignoring them as traditional models do [91]. This might result in more complex models, but the challenges resulting from glucose homeostasis mentioned above justify their use for diabetes treatment. Third, machine learning approaches imply the risk of developing a model which can perform perfectly on the training data without generalizing well to unseen data, called over-fitting [91]. It is therefore important to correct for over-fitting by using cross-validation [92] and regularization techniques [91].

The first attempt toward including machine learning algorithms into diabetes care was the use of supervised learning with artificial neural network (ANN) classification for diabetes treatment [93]. ANN algorithms infer a function minimizing the error between calculated parameters and desired parameters with the help of supervised/labeled training data. Of note, in supervised learning, the labeling of data needs to be performed by an expert, which is time consuming and prone to human error. In terms of blood glucose concentration prediction and insulin regimen recommendations, ANNs work well in short-term predictions [88] even in closed-loop systems [94]. However, they have not yet been tested for long-term predictions of the blood glucose concentration. Supervised learning systems, such as that used by Robertson *et al.*, need good training data which include the desired response [93]. These data can be expensive and/or time-consuming to collect, they assume that good responses are known, and in general only lead to reliable predictions for situations similar to those in the training data [92]. Furthermore, glucose control in diabetes patients is an ongoing task requiring regular control responses. The static input-output nature of supervised learning ignores this and therefore errors might propagate. An alternative approach is to describe the metabolic system as a process, with more desirable and less desirable metabolic states, and which responds to insulin doses (and other activities) by changing between these states. Learning directly on such a process can avoid the need for these expensive error-prone static labels and improve the coverage of solutions. To find a good control strategy for this process, we need an adaptive machine learning approach that can be trained directly on such a model, and this suggests the use of RL algorithms.

The need for reinforcement learning in glucose regulation

RL is a branch of machine learning, concerned with how an agent chooses actions to control a system. It is suited to problems including sequences of decisions along a timeline. Additionally, it can be used when decisions depend on the observed state, where effects may be remote in time from actions that induce them, and where there is some notion of preferred state(s) for the system. This is true for the artificial pancreas system, as there is a need to continuously observe the patients glucose concentration and determine the ideal time and amount for insulin delivery. Moreover, RL can be performed directly on real data, or it can interact with a dynamical system represented by a mathematical model and in general it makes very modest assumptions about this system [95]. A broad classification of the types of control algorithms in terms of performance (glucose control, amount of delivered insulin, reaction time of the system) or personalization is shown in Figure 4.

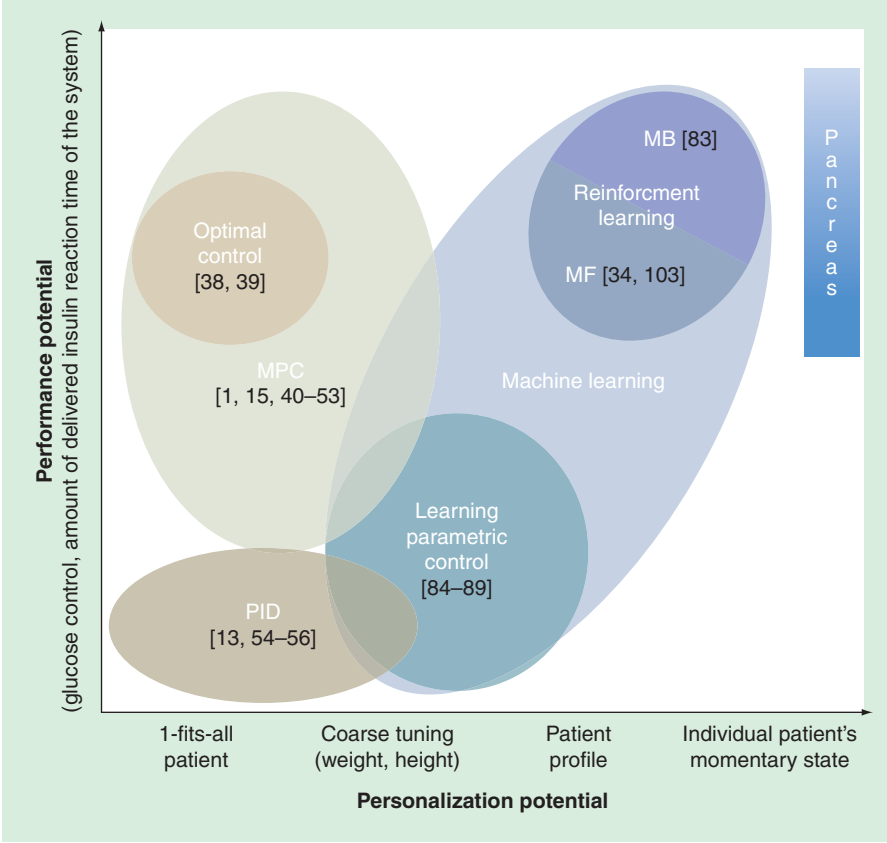

Figure 4. Commonly used types of algorithms for glucose control as well as machine learning and reinforcement learning algorithms. The y-axis reflects qualitative dimensions along which we expect performance (tracking target values, while minimizing dosage and maximizing desirable effects) to reside, whereas the x-axis reflects dimensions for increased personalization potential. The integration is based on our general assessment of the type of algorithms used (colored 'clouds' reflecting control engineering derived algorithms [earth colors] and machine learning derived algorithms [sea colors]). The references inside the clouds are to be seen as illustrative examples without ranking (color figure can be found online at: www.expert-reviews.com/doi/full/10.1586/17434440.2013.827515). MB: Model-based; MF: Model-free; MPC: Model predictive control; PID: Proportional integral derivative control.

The principle of RL is based on the interaction between a decision-making and self-learning agent and its environment. At each time point, the agent chooses an action to modify the environment. The environment changes its state and sends this information and a numerical reward according to the previous action back to the agent. Mapping of a particular state to a certain action is called policy of the algorithm and defines the behavior of the agent at each time step [95]. The goal of RL is to learn an optimal policy and thus maximize the amount of reward it receives over time [95]. To achieve that goal, the agent should not only choose the action which brings the most reward in one run (exploitation), it should also consider other possibilities to increase the overall reward (exploration). A balancing between exploitation and exploration is needed to generalize from experience. The agent need to explore unusual states in the system by occasionally choosing unpromising actions during the learning procedure, in order to choose good actions for even these rare states during normal operation. The procedure of a RL algorithm for diabetes is depicted in Figure 5.

In comparison with other traditional control strategies, RL does not require a detailed description of the environment in terms of a well-represented model [96] or labeled training data as in supervised learning strategies. After a learning procedure, the agent develops a policy and thus a control strategy from experience to predict certain situations and rewards without a necessary mathematical specification of the environment. Another advantage of RL algorithms is that they are uniquely suited to systems with inherent time delays [97] as these are present due to the subcutaneous glucose measurements and insulin injections. RL can also be used with large or even infinite state sets, which makes that approach useful for the different glycemic concentrations that occur during continuous glucose measuring.

Two potential criticisms of conventional RL methods are relevant to the case of insulin delivery control: The learned control is black-box, meaning it cannot be readily reused or generalized from, and they are not very efficient in terms of data. The efficacy issue arises, because conventional RL algorithms do not build explicit models of the environment, and are therefore sometimes referred to as model-free RL. To address this, we recommend the use of one or both of the following techniques: model-based or data-efficient RL. Model-based RL builds a dynamical model of the control problem through experience and uses this model to train an on-board model-free RL algorithm, for example, Dyna-Q [98]. Here, we define a model-based RL algorithm as one that maintains a system model, which is updated on-line, for example, from real data as it is observed, and which optimizes a reinforcement learner using this model. This model can be either entirely constructed from empirical data, or can use prior knowledge to constrain the family of models considered. For an insulin delivery system, a model-based reinforcement learner could therefore define *a priori* a dynamical system structure where blood glucose depends on insulin and beta-cells in a specified way, but use real data to update parameters of that model. Data-efficient RL algorithms focus on making the most efficient use of experience gained so far, for example, fitted-Q [99]. The latter approach has been proposed for use in clinical

domains [96,100], and recent work, which combines this with a model-based RL approach, has shown remarkable data efficiency in robotics tasks [101].

To date, RL algorithms have been proposed for the treatment of epilepsy [96], renal anemia [102] or the control of anesthesia [97]. In case of renal anemia, the RL algorithm was informed by an MPC, showing that these two approaches do not necessarily exclude each other but can be used in parallel. When used in closed-loop control of anesthesia *in silico*, a RL algorithm outperforms a PID control algorithm by less overshoot of the depth of hypnosis and faster achievement of steady state [103]. Thus, use of the RL algorithm in this closed-loop control setting resulted in tighter control, a principle that could be administrable for patients with diabetes. An initial study on using a RL algorithm to control an artificial pancreas reported good performance in controlling hyperglycemia *in silico* [25]. In this study, the state was defined as different glycemic ranges, action was defined as insulin infusion and reward was set equal to the difference of the glucose concentration from its target value [25]. *In silico* application of the algorithm led to correction of hyperglycemia to normoglycemia. However, given the criticism on RL mentioned above it is important to point out that this research used purely model-free RL in an off-line manner on a fixed model with fixed parameters, which is distinct from a model-based reinforcement learner that updates its model on-line (and typically consists of an internal adaptive model coupled with the more conventional model-free RL). There is no description of how to verify the accuracy of the model or how to adapt the controller to individual patients, although the authors do acknowledge that these are research issues. Furthermore, the tests included *in silico* patients only and were especially not yet performed *in vivo*. The same is true for an actor-critic control algorithm inspired by the principles of RL, which showed promising results in adults and children *in silico*, but still has to be verified in a clinical trial [104]. Moreover, the authors do not discuss how the early exploratory phase of the algorithm can be safely achieved *in vivo*. Thus, only a first proof of principle for the usability of RL algorithms in diabetes has been successfully performed. It would be worth comparing such RL algorithms with other algorithms in a larger setting in future studies and to further exploit their full potential in terms of flexible reactions on changes in the blood glucose concentrations of diabetes patients.

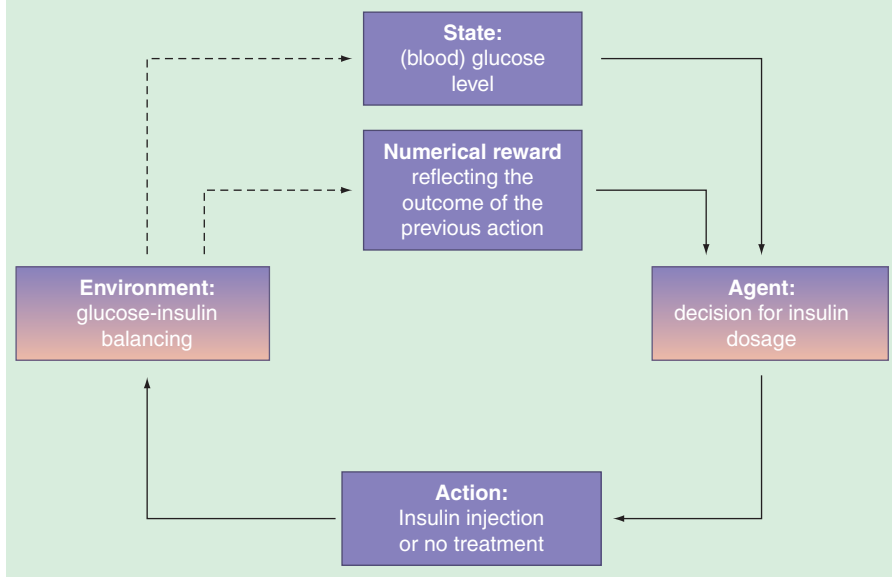

Figure 5. Reinforcement learning algorithms for diabetes. Changes in the state lead to an action of the agent, which changes the environment. The agent receives a numerical reward from the environment, which together with the next status will influence the next action.

Expert commentary

Algorithms for closed-loop models of insulin treatment have to deal with demanding challenges due to the complex physiology of glucose homeostasis as well as technical limitations of the components of an artificial pancreas. The flexible reactivity especially required for outpatient use suggests the use of data-driven machine learning algorithms. Among those, RL algorithms exhibit a great potential to deal with the time delay produced by the CGM system. In view of the available evidence, it can be summarized that RL algorithms provide a very promising approach for flexibly and independently maintaining normoglycemia in artificial pancreas systems. To date, the vast majority of papers reporting the development of algorithms demonstrate control in limited scenarios *in silico*, for example, an insulin spike after a single meal. We assert that stochastic models are essential to assess the reliability and stability of an algorithm for periods containing multiple meal events, whereas future *in vivo* studies of closed-loop algorithms are required to reliably assess performance and personalization.

Five-year view

In the future, minimization of the time delay between changes in the glucose concentration and the full effect of insulin as well as maximization of the accuracy of the subcutaneous glucose measuring devices will be the subject of studies on the components of artificial pancreas. Ultra-rapid acting insulins are under development as well as substances like GLP-1 or amylin, increasing the applicability of these approaches to glucose regulation. However, the more other substances for regulation of blood glucose come up, the more individual and flexible becomes glucose control, and the more complex. This increases the need for smart, personalized algorithms calculating insulin delivery in the future and might among others be achieved by the use of RL algorithms. In the future, closed-

loop drug delivery such as here in the case of insulin could be complemented with context-aware info-services (as developed on smartphones) to tune closed-loop drug delivery systems to the patient's momentary state.

Currently, control system and adaptive control system approaches to insulin delivery are based on systems that can only be adaptive up to the point of delivery to the patient, due to regulatory demands. However, true patient personalization requires the ability of the system to adapt to the changing daily routines of patient, for example, going on holiday or catching a stomach bug. Taking current regulatory trends into consideration it appears that model-based RL systems (MB-RL) have the best chance to be approved for adaptive personalization after delivery to the patients, as they can quantitatively predict the consequences of their actions and monitor their performance and effectiveness with respect to their own predictions and that of clinical references models to ensure patient safety.

Acknowledgements

MK Bothe wrote the manuscript. L Dickens, K Reichel, A Tellmann, B Ellger, M Westphal and AA Faisal reviewed and edited the manuscript.

Financial & competing interests disclosure

M Westphal, A Tellmann, K Reichel and M Bothe are employees of Fresenius Kabi Deutschland GmbH, Germany. Fresenius Kabi Deutschland GmbH distributes pumps for insulin delivery. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Key issues

- Maintaining normoglycemia is crucial in patients with diabetes mellitus or severe illness and is usually achieved by administration of insulin.
- An artificial pancreas system for closed-loop insulin delivery consists of a continuous glucose monitoring device, an algorithm calculating the correct amount of insulin and a pump delivering insulin.
- Current challenges for calculation of the correct dose include technical issues most notably with regard to the time delay between changes in the glucose concentration and the maximum effect of insulin.
- Use of newly identified substances for glucose regulation, such as glucagon or amylin, increases both the required flexibility and the complexity of the approach.
- Have reliable and safe coverage for all feasible metabolic states and to respond appropriately in novel situations.
- The individualized treatment regimes and the complex glucose regulating parameters elevate the need for smart algorithms that have reliable and safe coverage for all feasible metabolic states and respond appropriately in novel situations.
- Algorithms used in the past were initially based on model predictive control or proportional integral derivative control.
- Machine learning algorithms and especially reinforcement learning algorithms provide the advantages to learn the individual glucose pattern of a diabetic patient in spite of a time delay and to handle complex and external information to provide adaptive drug delivery after a learning procedure.
- For machine learning approaches, care must be taken to acquire appropriate data for the learning phase whereby the data should be representative, sufficient and optimally noise-reduced.
- To maximize the effectiveness of data-driven approaches, cross-validation and regularization techniques should be used and an extensive testing phase has to be performed.
- Glucose regulation is more than just beta-cell dynamics, therefore we need smarter algorithms that learn to take into account the bigger picture.

References

Papers of special note have been highlighted as:
• of interest
•• of considerable interest

1 Pachler C, Plank J, Weinhandl H *et al.* Tight glycaemic control by an automated algorithm with time-variant sampling in medical ICU patients. *Intensive Care Med.* 34(7), 1224–1230 (2008).

2 Elleri D, Allen JM, Biagioni M *et al.* Evaluation of a portable ambulatory prototype for automated overnight closed-loop insulin delivery in young people with type 1 diabetes. *Pediatr. Diab.* 13(6), 449–453 (2012).

3 Atlas E, Nimri R, Miller S, Grunberg EA, Phillip M. MD-logic artificial pancreas system: a pilot study in adults with type 1 diabetes. *Diab. Care* 33(5), 1072–1076 (2010).

4 International Diabetes Federation. IDF Diabetes Atlas (5th edition). Brussels, Belgium (2011)

5 Clemens AH, Chang PH, Myers RW. The development of Biostator, a Glucose Controlled Insulin Infusion System (GCIIS). *Horm. Metab. Res.* (Suppl. 7), 23–33 (1977).

6 American Diabetes Association. Standards of medical care in diabetes – 2011. *Diab. Care* 34(Suppl. 1), S11–S61 (2011).

7 Vaddiraju S, Burgess DJ, Tomazos I, Jain FC, Papadimitrakopoulos F. Technologies for continuous glucose monitoring: current problems and future promises. *J. Diab. Sci. Technol.* 4(6), 1540–1562 (2010).

8 Clinical and Laboratory Standards Institute. Performance Metrics for Continuous

Intersititial Glucose Monitoring; Approved Guideline. Clinical and Laboratory Standards Institute document POCT05-A. Wayne: Clinical and Laboratory Standards Institute; (2008).

9 Cengiz E, Tamborlane WV. A tale of two compartments: interstitial versus blood glucose monitoring. *Diab. Technol. Ther.* 11(Suppl. 1), S11–S16 (2009).

10 Aussedat B, Dupire-Angel M, Gifford R, Klein JC, Wilson GS, Reach G. Interstitial glucose concentration and glycemia: implications for continuous subcutaneous glucose monitoring. *Am. J. Physiol. Endocrinol. Metab.* 278(4), E716–E728 (2000).

11 Lonnroth P, Jansson PA, Smith U. A microdialysis method allowing characterization of intercellular water space in humans. *Am. J. Physiol.* 253(2 Pt 1), E228–E231 (1987).

12 Rebrin K, Steil GM, van Antwerp WP, Mastrototaro JJ. Subcutaneous glucose predicts plasma glucose independent of insulin: implications for continuous monitoring. *Am. J. Physiol.* 277(3 Pt 1), E561–E571 (1999).

13 Baek YH, Jin HY, Lee KA *et al.* The Correlation and accuracy of glucose levels between interstitial fluid and venous plasma by continuous glucose monitoring system. *Korean Diab. J.* 34(6), 350–358 (2010).

14 Kulcu E, Tamada JA, Reach G, Potts RO, Lesho MJ. Physiological differences between interstitial glucose and blood glucose measured in human subjects. *Diab. Care* 26(8), 2405–2409 (2003).

15 Beier B, Musick K, Matsumoto A, Panitch A, Nauman E, Irazoqui P. Toward a continuous intravascular glucose monitoring system. *Sensors* 11(1), 409–424 (2011).

16 Skjaervold NK, Solligard E, Hjelme DR, Aadahl P. Continuous measurement of blood glucose: validation of a new intravascular sensor. *Anesthesiology* 114(1), 120–125 (2011).

17 Aye T, Block J, Buckingham B. Toward closing the loop: an update on insulin pumps and continuous glucose monitoring systems. *Endocrinol. Metab. Clin. North America* 39(3), 609–624 (2010).

18 Keenan DB, Mastrototaro JJ, Voskanyan G, Steil GM. Delays in minimally invasive continuous glucose monitoring devices: a review of current technology. *J. Diab. Sci. Technol.* 3(5), 1207–1214 (2009).

19 Rebrin K, Sheppard NF Jr, Steil GM. Use of subcutaneous interstitial fluid glucose to estimate blood glucose: revisiting delay and sensor offset. *J. Diab. Sci. Technol.* 4(5), 1087–1098 (2010).

20 Mastrototaro J. The MiniMed Continuous Glucose Monitoring System (CGMS). *J. Pediatr. Endocrinol. Metab.* 12(Suppl. 3), 751–758 (1999).

21 Glucose Monitoring Study Group. Continuous glucose monitoring by means of the microdialysis technique: underlying fundamental aspects. *Diab. Technol. Ther.* 5(4), 545–561 (2003).

22 Banting FG, Campbell WR, Fletcher AA. Further clinical experience with insulin (pancreatic extracts) in the treatment of diabetes mellitus. *Br. Med. J.* 1(3236), 8–12 (1923).

23 Gilchrist JA, Best CH, Banting FG. Observations with insulin on department of soldiers' civil Re-establishment diabetics. *Canadian Med. Assoc. J.* 13(8), 565–572 (1923).

24 Renard E. Insulin delivery route for the artificial pancreas: subcutaneous, intraperitoneal, or intravenous? *Pros and cons. J. Diab. Sci. Technol.* 2(4), 735–738 (2008).

25 Yasini S, Naghibi-Sistani MB, Karimpour A. Agent-based Simulation for Blood Glucose Control in Diabetic Patients. *World Acad. Sci. Eng. Technol.* 57(33), 11 (2009).

26 McMahon GT, Arky RA. Inhaled insulin for diabetes mellitus. *N. Engl. J. Med.* 356(5), 497–502 (2007).

27 Farmer TG, Jr., Edgar TF, Peppas NA. The future of open- and closed-loop insulin delivery systems. *J. Pharm. Pharmacol.* 60(1), 1–13 (2008).

28 Kapitza C, Zijlstra E, Heinemann L, Castelli MC, Riley G, Heise T. Oral insulin: a comparison with subcutaneous regular human insulin in patients with type 2 diabetes. *Diab. Care* 33(6), 1288–1290 (2010).

29 Najafzadeh H, Kooshapur H, Kianidehkordi F. Evaluation of an oral insulin formulation in normal and diabetic rats. *Indian J. Pharm.* 44(1), 103–105 (2012).

30 Diabetes Control and Complications Trial Research Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. The Diabetes Control and Complications Trial Research Group. *N. Engl. J. Med.* 329(14), 977–986 (1993).

31 Kumareswaran K, Evans ML, Hovorka R. Closed-loop insulin delivery: towards improved diabetes care. *Discov. Med.* 13(69), 159–170 (2012).

32 Thorens B. Brain glucose sensing and neural regulation of insulin and glucagon secretion. *Diab. Obesity Metab.* 13(Suppl. 1), 82–88 (2011).

33 Thorens B. Sensing of glucose in the brain. *Handb. Exp. Pharmacol.* (209), 277–294 (2012).

34 Weinzimer SA, Steil GM, Swan KL, Dziura J, Kurtz N, Tamborlane WV. Fully automated closed-loop insulin delivery versus semiautomated hybrid control in pediatric patients with type 1 diabetes using an artificial pancreas. *Diab. Care* 31(5), 934–939 (2008).

35 Renard E, Place J, Cantwell M, Chevassus H, Palerm CC. Closed-loop insulin delivery using a subcutaneous glucose sensor and intraperitoneal insulin delivery: feasibility study testing a new model for the artificial pancreas. *Diab. Care* 33(1), 121–127 (2010).

36 Dassau E, Bequette BW, Buckingham BA, Doyle FJ 3rd. Detection of a meal using continuous glucose monitoring: implications for an artificial beta-cell. *Diab. Care* 31(2), 295–300 (2008).

37 Lee H, Buckingham BA, Wilson DM, Bequette BW. A closed-loop artificial pancreas using model predictive control and a sliding meal size estimator. *J. Diab. Sci. Technol.* 3(5), 1082–1090 (2009).

38 Panteleon AE, Loutseiko M, Steil GM, Rebrin K. Evaluation of the effect of gain on the meal response of an automated closed-loop insulin delivery system. *Diabetes* 55(7), 1995–2000 (2006).

39 Steil GM, Palerm CC, Kurtz N *et al.* The effect of insulin feedback on closed loop glucose control. *J. Clin. Endocrinol. Metab.* 96(5), 1402–1408 (2011).

40 Wintergerst KA, Deiss D, Buckingham B *et al.* Glucose control in pediatric intensive care unit patients using an insulin-glucose algorithm. *Diab. Technol. Ther.* 9(3), 211–222 (2007).

41 Acikgoz US, Diwekar UM. Blood glucose regulation with stochastic optimal control for insulin-dependent diabetic patients. *Chem. Eng. Sci.* 65(3), 1227–1236 (2010).

42 Chee F, Savkin AV, Fernando TL, Nahavandi S. Optimal H infinity insulin injection control for blood glucose regulation in diabetic patients. *IEEE Trans. Biomed. Eng.* 52(10), 1625-–1631 (2005).

43 Bruttomesso D, Farret A, Costa S *et al.* Closed-loop artificial pancreas using subcutaneous glucose sensing and insulin

delivery and a model predictive control algorithm: preliminary studies in Padova and Montpellier. *J. Diab. Sci. Technol.* 3(5), 1014–1021 (2009).

44 Clarke WL, Anderson S, Breton M, Patek S, Kashmer L, Kovatchev B. Closed-loop artificial pancreas using subcutaneous glucose sensing and insulin delivery and a model predictive control algorithm: the Virginia experience. *J. Diab. Sci. Technol.* 3(5), 1031–1038 (2009).

45 Elleri D, Allen JM, Nodale M *et al.* Automated overnight closed-loop glucose control in young children with type 1 diabetes. *Diab. Technol. Ther.* 13(4), 419–424 (2011).

46 Ellingsen C, Dassau E, Zisser H *et al.* Safety constraints in an artificial pancreatic beta cell: an implementation of model predictive control with insulin on board. *J. Diab. Sci. Technol.* 3(3), 536–544 (2009).

47 Gillis R, Palerm CC, Zisser H, Jovanovic L, Seborg DE, Doyle FJ. Glucose estimation and prediction through meal responses using ambulatory subject data for advisory mode model predictive control. *J. Diab. Sci. Technol.* 1(6), 825–833 (2007).

48 Grosman B, Dassau E, Zisser HC, Jovanovic L, Doyle FJ 3rd. Zone model predictive control: a strategy to minimize hyper- and hypoglycemic events. *J. Diab. Sci. Technol.* 4(4), 961–975 (2010).

49 Hovorka R, Canonico V, Chassin LJ *et al.* Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes. *Physiol. Meas.* 25(4), 905–920 (2004).

50 Kan S, Onodera H, Furutani E *et al.* Novel control system for blood glucose using a model predictive method. *ASAIOJ.* 46(6), 657–662 (2000).

51 Magni L, Forgione M, Toffanin C *et al.* Run-to-run tuning of model predictive control for type 1 diabetes subjects: in silico trial. *J. Diab. Sci. Technol.* 3(5), 1091–1098 (2009).

52 Murphy HR, Elleri D, Allen JM *et al.* Closed-loop insulin delivery during pregnancy complicated by type 1 diabetes. *Diab. Care* 34(2), 406–411 (2011).

53 Schaller HC, Schaupp L, Bodenlenz M *et al.* On-line adaptive algorithm with glucose prediction capacity for subcutaneous closed loop control of glucose: evaluation under fasting conditions in patients with Type 1 diabetes. *Diab. Med.* 23(1), 90–93 (2006).

54 Schlotthauer G, Gamero LG, Torres ME, Nicolini GA. Modeling, identification and nonlinear model predictive control of type I diabetic patient. *Med. Eng. Phys.* 28(3), 240–250 (2006).

55 Wang Y, Dassau E, Doyle FJ 3rd. Closed-loop control of artificial pancreatic Beta-cell in type 1 diabetes mellitus using model predictive iterative learning control. *IEEE trans. Biomed. Eng.* 57(2), 211–219 (2010).

56 Wilinska ME, Budiman ES, Taub MB *et al.* Overnight closed-loop insulin delivery with model predictive control: assessment of hypoglycemia and hyperglycemia risk using simulation studies. *J. Diab. Sci. Technol.* 3(5), 1109–1120 (2009).

57 Elleri D, Dunger DB, Hovorka R. Closed-loop insulin delivery for treatment of type 1 diabetes. *BMC Med.* 9, 120 (2011).

58 Hovorka R, Kumareswaran K, Harris J *et al.* Overnight closed loop insulin delivery (artificial pancreas) in adults with type 1 diabetes: crossover randomised controlled studies. *BMJ* 342, d1855 (2011).

59 Nimri R, Atlas E, Ajzensztejn M, Miller S, Oron T, Phillip M. Feasibility study of automated overnight closed-loop glucose control under MD-logic artificial pancreas in patients with type 1 diabetes: the DREAM Project. *Diab. Technol. Ther.* 14(8), 728–735 (2012).

60 Hovorka R, Kremen J, Blaha J *et al.* Blood glucose control by a model predictive control algorithm with variable sampling rate versus a routine glucose management protocol in cardiac surgery patients: a randomized controlled trial. *J. Clin. Endocrinol. Metab.* 92(8), 2960–2964 (2007).

61 Egi M, Finfer S, Bellomo R. Glycemic control in the ICU. *Chest* 140(1), 212–220 (2011).

62 Kavanagh BP. Glucose in the ICU - evidence, guidelines, and outcomes. *N. Engl. J. Med.* 367(13), 1259–1260 (2012).

63 Cordingley JJ, Vlasselaers D, Dormand NC *et al.* Intensive insulin therapy: enhanced Model Predictive Control algorithm versus standard care. *Intensive Care Med.* 35(1), 123–128 (2009).

64 Blaha J, Kopecky P, Matias M *et al.* Comparison of three protocols for tight glycemic control in cardiac surgery patients. *Diab. Care* 32(5), 757–761 (2009).

65 Breton M, Farret A, Bruttomesso D *et al.* Fully integrated artificial pancreas in type 1 diabetes: modular closed-loop glucose control maintains near normoglycemia. *Diabetes* 61(9), 2230–2237 (2012).

66 Cobelli C, Renard E, Kovatchev BP *et al.* Pilot studies of wearable outpatient artificial pancreas in type 1 diabetes. *Diab. Care* 35(9), e65–e67 (2012).

67 Grayson BE, Seeley RJ, Sandoval DA. Wired on sugar: the role of the CNS in the regulation of glucose homeostasis. *Nat. Rev. Neurosci.* 14(1), 24–37 (2013).

68 Teixeira RE, Malin S. The next generation of artificial pancreas control algorithms. *J. Diab. Sci. Technol.* 2(1), 105–112 (2008).

69 D'Alessio DA, Vahl TP. Glucagon-like peptide 1: evolution of an incretin into a treatment for diabetes. *Am. J. Physiol. Endocrinol. Metab.* 286(6), E882–E890 (2004).

70 Bergman RN, Phillips LS, Cobelli C. Physiologic evaluation of factors controlling glucose tolerance in man: measurement of insulin sensitivity and beta-cell glucose sensitivity from the response to intravenous glucose. *J. Clin. Invest.* 68(6), 1456–1467 (1981).

71 Roy A, Parker RS. Mixed meal modeling and disturbance rejection in type I diabetic patients. *Conf. Proc.* 1, 323–326 (2006).

72 Steil GM, Panteleon AE, Rebrin K. Closed-loop insulin delivery-the path to physiological glucose control. *Adv. Drug Deliv. Rev.* 56(2), 125–144 (2004).

73 Luzi L, DeFronzo RA. Effect of loss of first-phase insulin secretion on hepatic glucose production and tissue glucose disposal in humans. *Am. J. Physiol.* 257(2 Pt 1), E241–E246 (1989).

74 Hartman I. Insulin analogs: impact on treatment success, satisfaction, quality of life, and adherence. *Clin. Med. Res.* 6(2), 54–67 (2008).

75 Castle JR, Engle JM, El Youssef J *et al.* Novel use of glucagon in a closed-loop system for prevention of hypoglycemia in type 1 diabetes. *Diab. Care* 33(6), 1282–1287 (2010).

76 El-Khatib FH, Russell SJ, Nathan DM, Sutherlin RG, Damiano ER. A bihormonal closed-loop artificial pancreas for type 1 diabetes. *Sci. Trans. Med.* 2(27), 27ra27 (2010).

77 Adeghate E, Kalasz H. Amylin analogues in the treatment of diabetes mellitus: medicinal chemistry and structural basis of its function. *Open Med. Chem. J.* 5(Suppl. 2), 78–81 (2011).

78 Schmitz O, Brock B, Rungby J. Amylin agonists: a novel approach in the treatment of diabetes. *Diabetes* 53(Suppl. 3), S233–S238 (2004).

79 Heptulla RA, Rodriguez LM, Mason KJ, Haymond MW. Twenty-four-hour simultaneous subcutaneous Basal-bolus administration of insulin and amylin in adolescents with type 1 diabetes decreases postprandial hyperglycemia. *J. Clin. Endocrinol. Metab*. 94(5), 1608–1611 (2009).

80 Weinzimer SA, Sherr JL, Cengiz E *et al.* Effect of pramlintide on prandial glycemic excursions during closed-loop control in adolescents and young adults with type 1 diabetes. *Diab. Care* 35(10), 1994–1999 (2012).

81 Macconell L, Brown C, Gurney K, Han J. Safety and tolerability of exenatide twice daily in patients with type 2 diabetes: integrated analysis of 5594 patients from 19 placebo-controlled and comparator-controlled clinical trials. *Diab. Metab. Syndr. Obes*. 5, 29–41 (2012).

82 Nikfar S, Abdollahi M, Salari P. The efficacy and tolerability of exenatide in comparison to placebo; a systematic review and meta-analysis of randomized clinical trials. *J. Pharm. Pharm. Sci*. 15(1), 1–30 (2012).

83 US Food and Drug Administration (FDA) - Center for Drug Evaluation and Research. Reports of altered kidney function in patients using exenatide (marketed as Byetta). Information for Healthcare Professionals. Marketed as Byetta, 1, (2009). www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/DrugSafetyInformationforHeathcareProfessionals/ucm188656.htm (Accessed 1 August 2013).

84 Reichel K, Dickens L, Tellmann A, Bothe MK, Westphal M, Faisal AA. Machine Learning for Closed-loop Insulin Delivery: Model & in-vivo Studies. *Conference: Bioengineering* Oxford, UK, 12, (2012).

85 Dazzi D, Taddei F, Gavarini A, Uggeri E, Negro R, Pezzarossa A. The control of blood glucose in the critical diabetic patient: a neuro-fuzzy method. *J. Diab. Complications* 15(2), 80–87 (2001).

86 El-Jabali AK. Neural network modeling and control of type 1 diabetes mellitus. *Bioprocess Biosys. Eng*. 27(2), 75–79 (2005).

87 Gogou G, Maglaveras N, Ambrosiadou BV, Goulis D, Pappas C. A neural network approach in diabetes management by insulin administration. *J. Med. Syst*. 25(2), 119–131 (2001).

88 Mougiakakou SG, Nikita KS. A neural network approach for insulin regime and dose adjustment in type 1 diabetes. *Diab. Technol. Ther*. 2(3), 381–389 (2000).

89 Otto E, Semotok C, Andrysek J, Basir O. An intelligent diabetes software prototype: predicting blood glucose levels and recommending regimen changes. *Diab. Technol. Ther*. 2(4), 569–576 (2000).

90 Trajanoski Z, Regittnig W, Wach P. Simulation studies on neural predictive control of glucose using the subcutaneous route. *Comput. Methods Prog. Biomed*. 56(2), 133–139 (1998).

91 Bishop CM. Pattern Recognition and Machine Learning (Information Science and Statistics). Springer-Verlag New York, Inc., Secaucus, NJ (2006).

92 Waljee AK, Higgins PD. Machine learning in medicine: a primer for physicians. *Am. J. Gastroenterol*. 105(6), 1224–1226 (2010).

93 Robertson G, Lehmann ED, Sandham W, Hamilton D. Blood glucose prediction using artificial neural networks trained with the aida diabetes simulator: a proof-of-concept pilot study. *J. Electrical Comput. Eng*. 2011, ID 681786 (2011).

94 Zarkogianni K, Vazeou A, Mougiakakou SG, Prountzou A, Nikita KS. An insulin infusion advisory system based on autotuning nonlinear model-predictive control. *IEEE Trans. Biomed. Eng*. 58(9), 2467–2477 (2011).

95 Sutton RS, Barto AG. Reinforcement Learning: An Introduction. MIT press, Cambridge, 1(1), (1998).

96 Pineau J, Guez A, Vincent R, Panuccio G, Avoli M. Treating epilepsy via adaptive neurostimulation: a reinforcement learning approach. *Int. J. Neur. Sys*. 19(4), 227–240 (2009).

97 Moore BL, Doufas AG, Pyeatt LD. Reinforcement learning: a novel method for optimal control of propofol-induced hypnosis. *Anesth. Analg*. 112(2), 360–367 (2011).

98 Sutton RS. Integrated Architectures for Learning, Planning, and Reacting Based on Approximating Dynamic Programming. In *ML* 216–224 (1990).

99 Ernst D, Geurts P, Wehenkel L. Tree-based batch mode reinforcement learning. *J. Mach. Learn. Res*. 6, 503–556 (2005).

100 Ernst D, Stan G-B, Goncalves J, Wehenkel L. Clinical Data Based Optimal STI Strategies for HIV; A Reinforcement Learning Approach. In *Decision and Control, 2006 45th IEEE Conference on* 667–672, IEEE (2006).

101 Deisenroth MP, Rasmussen CE. PILCO: A Model-Based and Data-Efficient Approach to Policy Search. In *Proceedings of the 28th International Conference on Machine Learning, Bellevue, Washington, USA (ICML-11)* 465–472 (2011).

102 Gaweda AE, Muezzinoglu MK, Jacobs AA, Aronoff GR, Brier ME. Model predictive control with reinforcement learning for drug delivery in renal anemia management. *Conf. Proc*. 1, 5177–5180 (2006).

103 Moore BL, Quasny TM, Doufas AG. Reinforcement learning versus proportional-integral-derivative control of hypnosis in a simulated intraoperative patient. *Anesth. Analg*. 112(2), 350–359 (2011).

104 Daskalaki E, Diem P, Mougiakakou SG. An Actor-Critic based controller for glucose regulation in type 1 diabetes. *Comput. Methods Prog. Biomed*. 109(2), 116–125. (2013).